2012

Train Station, The Five-Minute Analyst

Schramm, Harrison

Analytics-Magazine.org, www.informs.org, Analytics, September/October 2012

http://hdl.handle.net/10945/40116

Train Station

How bad is it to get on the wrong train? We've carefully bounded the problem such that the only costs are in terms of time.

BY HARRISON SCHRAMM

Two men are traveling in a foreign city with no map and limited knowledge of the language, and, finished for the day, are trying to get back to their lodgings. They have arrived at a light rail station as a train is boarding. Should they board the train in front of them, not knowing if it is the "right" train, or should they wait for more information? Let's suppose that we have some time – five minutes, say – to analyze this decision.

Before proceeding, let's make some bounding assumptions. First, this is a city light rail, and for simplicity we will assume is similar to the Washington, D.C., Metro. If you miss the train, the next stop is only a few minutes away. Second, let's presume that the trains will be in operation longer than the time scope of the problem, so we are not concerned with being on the last train of the night. We also assume that you are able to change directions without leaving the train system. You took the train earlier in the day, but you didn't necessarily get off at the stop you are currently located. Finally, we assume that you will know with certainty as soon as the train leaves the station if you are headed in the correct direction. If we had no knowledge, we would think guessing will make the "right" choice 50 percent of the time.

How bad is it to get on the wrong train? We've carefully bounded the problem such that the only costs are in terms of time. Consider the paths of two travelers who we call Rob and Jeff, currently at Crystal City wanting to head south, but are unsure if "Huntington" or "Mt. Vernon Square" is the correct direction. At noon, Rob gets on a train headed toward Mt. Vernon Square while Jeff waits for the next train to Huntington. At 12:01, Rob realizes he has chosen the wrong train. At 12:03, he gets off at Pentagon City, crosses over to the other platform, and at 12:05 boards a train headed back toward Crystal City. At 12:10, he arrives back at Crystal City in time to see Jeff get on the train and sit next to him. This is because the whole time Rob was on the "wrong" train, he was headed toward the "right" one. This turns out to be the key to thinking about this problem.

Let's make this story concrete with some mathematics. Let $T_{A \to B}$ be the transit time between train stations A and B. Let τ be the time until the next train headed from B to A leaves station B. Rob will not be penalized for choosing the "wrong train" so long as

$$T_{A \to B} < \tau$$

Now we just have to figure out what this expression means to our travelers! Let's proceed by assuming that these times are exponentially distributed, but I'd like to explain the assumption. I'm doing it for two equally important reasons:

1. It makes the mathematics tractable.
2. It is a better model for reality.

I'd like to explain (2) first: We know that the time of train arrivals is continuous and positive, i.e., we cannot get on trains that have already departed the station. This gives us two "easy" distributions to choose from [1]. However, we're assuming that we've ridden on the train once already, and so we may wish to estimate the waiting time based on a single sample. This is a very dodgy business, but we have no choice. If we say that train arrivals are uniformly distributed, then we are implicitly saying that we know the longest time we would ever have to wait. If this were the case, we would know the true upper bound on transit times, and the problem would be much simpler than what is posed here.

With regards to (1), we have to keep the math simple due to the five-minute analysis. There's a nice property of the exponential distribution that is worth remembering, and it is frequently called the "race of the exponentials." If T_A, T_B are exponentially distributed with parameters λ_A, λ_B, then:

$$\Pr\{T_A < T_B\} = \frac{\lambda_A}{\lambda_A + \lambda_B}.$$

We may use (2) to conclude that if the amount of time we waited for our train in the morning is approximately equal to the time between stations (not an unreasonable assumption), then if we get on the train before us, we will have made the correct choice with 75 percent probability.

A more sophisticated approach, beyond the scope of our current discussion, might use the

If your hotel is near the city centre and it is "early" when you are headed toward your hotel, you may presume that the majority of travelers at that time may be headed that way as well, and get on the train if the majority of your fellow platform-standers do as well.

number of persons at the station as an estimator for the time remaining until the next train. If we let N be the number of persons on the platform and T_R be the time remaining until the next train, we would be then interested in finding $E[T_R|N]$.

Bonus: There is another, simpler approach to making this decision that is also worth mentioning. If your hotel is near the city center and it is "early" [2] when you are headed toward your hotel, you may presume that the majority of travelers at that time may be headed that way as well, and get on the train if the majority of your fellow platform-standers do as well. ❙

Harrison Schramm *(harrison.schramm@gmail.com) is a military instructor in the Operations Research Department at the Naval Postgraduate School in Monterey, Calif. He is an INFORMS member.*

NOTES

1. This is not to say that we are restricted to these choices; it means that these are the ones that are attractive for this effort.
2. By which we mean "early" to the local custom. Americans may consider 5 p.m. as early, while Europeans may find anything before 11 p.m. as early.